new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US Attorney Docket No.: 2102680-990101 Gray Cary et al. – GTS/415-836-2500

Applicants: Victor Hsieh Sheet 1 of 36

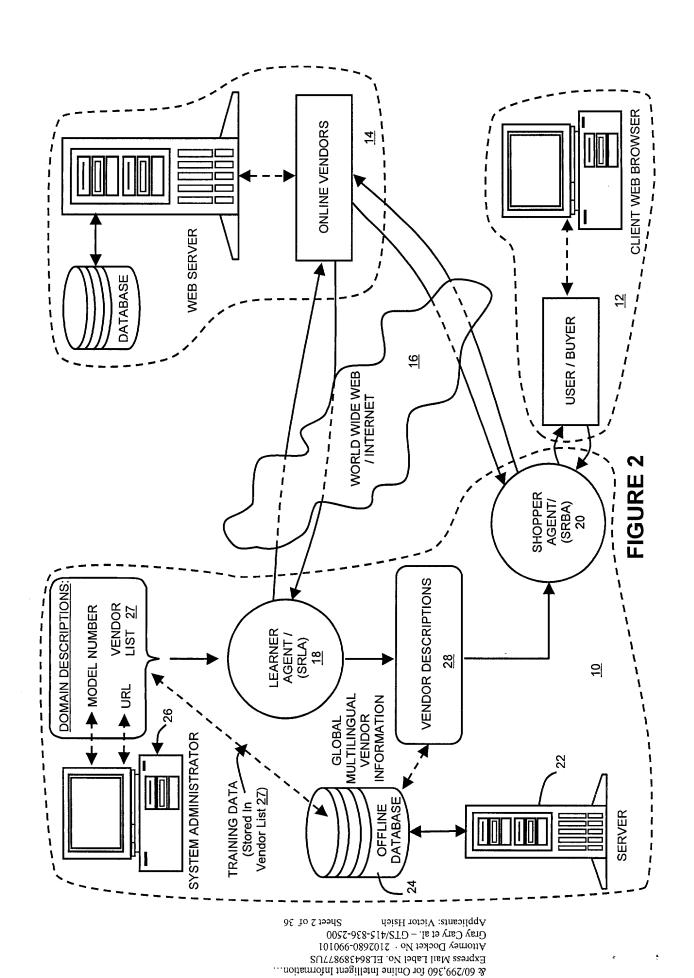
{7,","",""</I>,""</TABLE>"}

procedure pcwrapHLRT (page P)

scan to the line 7 in P

while the next occurrence of li in P occurs before the next occurrence of t scan in P to next occurrence of li; save position as start of item attribute scan in P to next occurrence of li; save position as end of item attribute scan in P to next occurrence of li; save position as start of price attribute scan in P to next occurrence of li; save position as end of price attribute return extracted , (item,price), } pairs

FIGURE 1 (PRIOR ART)



new U.S. Appln. based on Appln. Nos. 60/236,574

new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US Attorney Docket No.: 2102680-990101 Gray Cary et al. – GTS/415-836-2500 Applicants: Victor Hsieh Sheet 3 of 36

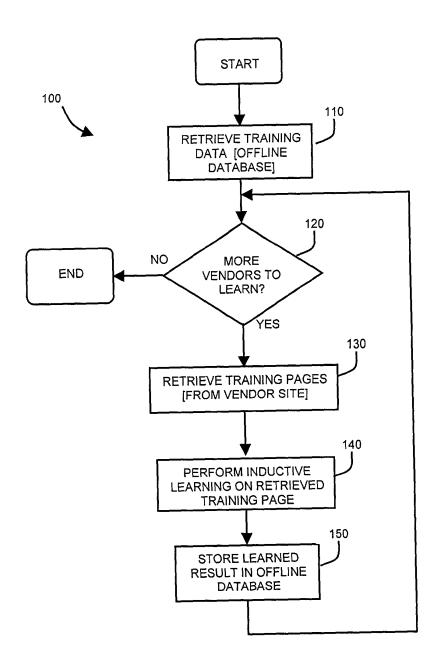


FIGURE 3

new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US
Attorney Docket No.: 2102680-990101
Gray Cary et al. – GTS/415-836-2500
Applicants Victor Hsieh Sheet 4 of 36

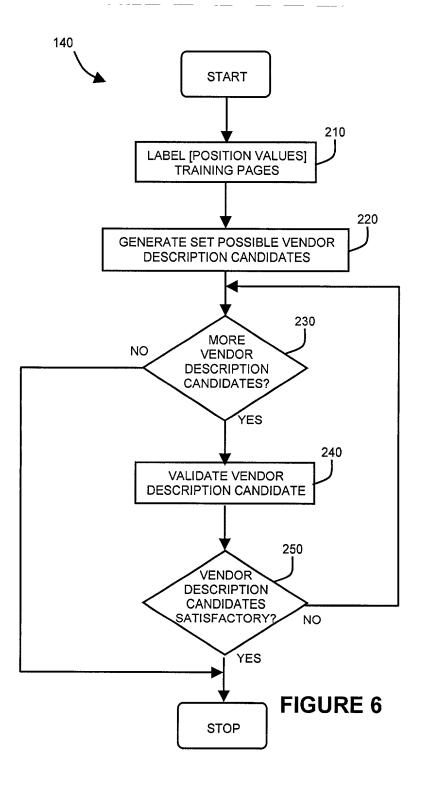
Element Name	Element Description
Vendor Name	Name of the online shop
Vendor URL	The URL of the online shop
Form URL	The URL to submit search data
Learning Domain	The domain used to learn the vendor description
Head	The end of header position of vendor's product pages. This element to learnt to reduce the product page searching space and thus shorten the product information extraction time
Tail	The start of footer position of vendor's product pages. Same as Head element, with this element the product information extraction time for Shopper Agent is shorten.
Left Delimiter of Item	The Shopper Agent uses this two delimiters to identify
Right Delimiter of Item	a product
Left Delimiter of Price	The Shopper Agent uses this two delimiters to locate
Right Delimiter of Price	a product's price information

FIGURE 4

	EXAMPLE
Element	Element Description
Name	·
Vendor Name	800.com
Vendor URL	http://www.800.com
Form URL	http://www.800.com/search/srchrslts.asp?qs=1&slteentry=All&entry=
Learning	Md
Domain	
Head	1230
Tail	
Left Delimiter	/td width="12">
of Item	
Right	
Delimiter of	
Item	
Left Delimiter	Your Price: \$
of Price	
Right	
Delimiter of	
Price	

FIGURE 5

new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US Attorney Docket No.: 2102680-990101 Gray Cary et al. – GTS/415-836-2500 Applicants: Victor Hsieh Sheet 5 of 36



new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US Attorney Docket No.: 2102680-990101 Gray Cary et al. – GTS/415-836-2500

Applicants: Victor Hsieh Sheet 6 of 3	Applicants: Victor Hsien Slice U UL 3
---------------------------------------	---------------------------------------

MD PRICE		<html></html>		
		<pre><title>A Simple Product Catalogs</title></pre>		
Model	PRICE	<body></body>		
Number	(US\$)	<pre><h2> MD Price </h2></pre>		
HM381MD	399.95	<pre><table border="1"> <tr bgcolor="ORANGE"><th> Model Number </th></tr></table></pre>	Model Number	
Model Number				
MD2070	599.95	<pre><th> PRICE(US\$)</th></pre>	PRICE(US\$)	
MD203	249.95	<pre><tr><td> HM381MD</td><td><i>399.95</i></td></tr></pre>	 HM381MD	<i>399.95</i>
 HM381MD	<i>399.95</i>			
MDR3	399.95	<tr><td> MD2070</td><td><i>599.95</i></td></tr>	 MD2070	<i>599.95</i>
 MD2070	<i>599.95</i>			
		<tr><td> MD203</td><td><i>249.95</i></td></tr>	 MD203	<i>249.95</i>
 MD203	<i>249.95</i>			
		<tr><td> MDR3</td><td><i>399.95</i></td></tr>	 MDR3	<i>399.95</i>
 MDR3	<i>399.95</i>			

		```  End Of The Product Catalog ```		

# FIGURE 7

Product Entry	Labels
{ HM381MD, 399.95}	{<<174, 180>, <197, 202>>}
{ MD2070, 599.95}	{<<229, 234>, <251, 256>>}
{ MD203, 249.95}	{<<283, 287>, <304, 309>>}
{ MDR3, 399.95}	{<<336, 339>, <356, 361>>}

# FIGURE 8

Product Entry	Labels
{ PRODUCT, PRICE}	{ <product delimiter,<br="" left="">PRODUCT RIGHT DELIMITER&gt;,</product>
	<pre><price delimiter="" delimiter,="" left="" price="" right="">}</price></pre>
* * *	* * *

# **FIGURE 9**

new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information.. Express Mail Label No. EL864389877US Attorney Docket No.: 2102680-990101 Gray Cary et al. – GTS/415-836-2500

Applicants: Victor Hsieh Sheet 7 of 36

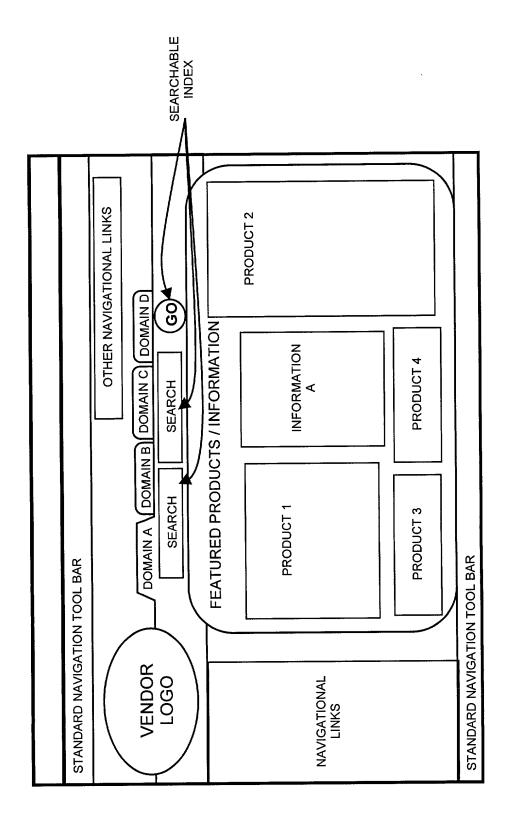
Delimiter	Candidates																									
Item Left Delimiter	<tr><td><b> TR&gt;<td><b> R&gt;<td><b>   <td><b>   <td><b>   <td><b>   <b>   <b> <b> <b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   &lt;</b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></td></b></td></b></td></b></td></b></td></b></td></tr> <tr><td>Item Right Delimiter</td><td></td><td><i></i></td><td><i< br=""></i<></td><td>&lt;                                                                    /B                       /B             /B          /B          /B       /B</td></tr> <tr><td>Price Left Delimiter</td><td></td><td>&lt; &gt; /B&gt;</td><td>&lt; &gt; B&gt;</td><td>&lt; &gt; &gt;&gt;TD&gt;<td>&lt; &gt; &gt;</td><td>&lt; &gt; </td><td>&lt; &gt; </td><td>&lt; &gt; /TD&gt;<td>&lt; &gt; /TD&gt;<td>&lt; &gt;  TD&gt;<td>&lt; &gt;  TD&gt;<td>&lt; &gt;  D&gt;<td>&lt; &gt;  TD&gt;&lt; &gt;  TD&gt;&lt; &gt; </td>&lt; &gt; </td>&lt; &gt; </td>&lt; &gt; </td>&lt; &gt;  &gt;&gt;TD&gt;&lt; &gt; </td>&lt; &gt;  &gt;&gt;TD&gt;&lt;   &gt;&gt;</td></td></tr>	<b> TR&gt;<td><b> R&gt;<td><b>   <td><b>   <td><b>   <td><b>   <b>   <b> <b> <b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   &lt;</b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></td></b></td></b></td></b></td></b></td></b>	<b> R&gt;<td><b>   <td><b>   <td><b>   <td><b>   <b>   <b> <b> <b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   &lt;</b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></td></b></td></b></td></b></td></b>	<b>   <td><b>   <td><b>   <td><b>   <b>   <b> <b> <b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   &lt;</b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></td></b></td></b></td></b>	<b>   <td><b>   <td><b>   <b>   <b> <b> <b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   &lt;</b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></td></b></td></b>	<b>   <td><b>   <b>   <b> <b> <b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   &lt;</b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></td></b>	<b>   <b>   <b> <b> <b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   &lt;</b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b>	Item Right Delimiter		<i></i>	<i< br=""></i<>	<                                                                    /B                       /B             /B          /B          /B       /B	Price Left Delimiter		< > /B>	< > B>	< > >>TD> <td>&lt; &gt; &gt;</td> <td>&lt; &gt; </td> <td>&lt; &gt; </td> <td>&lt; &gt; /TD&gt;<td>&lt; &gt; /TD&gt;<td>&lt; &gt;  TD&gt;<td>&lt; &gt;  TD&gt;<td>&lt; &gt;  D&gt;<td>&lt; &gt;  TD&gt;&lt; &gt;  TD&gt;&lt; &gt; </td>&lt; &gt; </td>&lt; &gt; </td>&lt; &gt; </td>&lt; &gt;  &gt;&gt;TD&gt;&lt; &gt; </td>&lt; &gt;  &gt;&gt;TD&gt;&lt;   &gt;&gt;</td>	< > >	< >	< >	< > /TD> <td>&lt; &gt; /TD&gt;<td>&lt; &gt;  TD&gt;<td>&lt; &gt;  TD&gt;<td>&lt; &gt;  D&gt;<td>&lt; &gt;  TD&gt;&lt; &gt;  TD&gt;&lt; &gt; </td>&lt; &gt; </td>&lt; &gt; </td>&lt; &gt; </td>&lt; &gt;  &gt;&gt;TD&gt;&lt; &gt; </td> < >  >>TD><   >>	< > /TD> <td>&lt; &gt;  TD&gt;<td>&lt; &gt;  TD&gt;<td>&lt; &gt;  D&gt;<td>&lt; &gt;  TD&gt;&lt; &gt;  TD&gt;&lt; &gt; </td>&lt; &gt; </td>&lt; &gt; </td>&lt; &gt; </td> < >  >>TD>< >	< >  TD> <td>&lt; &gt;  TD&gt;<td>&lt; &gt;  D&gt;<td>&lt; &gt;  TD&gt;&lt; &gt;  TD&gt;&lt; &gt; </td>&lt; &gt; </td>&lt; &gt; </td> < >	< >  TD> <td>&lt; &gt;  D&gt;<td>&lt; &gt;  TD&gt;&lt; &gt;  TD&gt;&lt; &gt; </td>&lt; &gt; </td> < >	< >  D> <td>&lt; &gt;  TD&gt;&lt; &gt;  TD&gt;&lt; &gt; </td> < >	< >  TD>< >  TD>< >
<b> TR&gt;<td><b> R&gt;<td><b>   <td><b>   <td><b>   <td><b>   <b>   <b> <b> <b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   &lt;</b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></td></b></td></b></td></b></td></b></td></b>	<b> R&gt;<td><b>   <td><b>   <td><b>   <td><b>   <b>   <b> <b> <b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   &lt;</b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></td></b></td></b></td></b></td></b>	<b>   <td><b>   <td><b>   <td><b>   <b>   <b> <b> <b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   &lt;</b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></td></b></td></b></td></b>	<b>   <td><b>   <td><b>   <b>   <b> <b> <b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   &lt;</b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></td></b></td></b>	<b>   <td><b>   <b>   <b> <b> <b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   &lt;</b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></td></b>	<b>   <b>   <b> <b> <b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   <b <br=""></b>   &lt;</b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b></b>																					
Item Right Delimiter		<i></i>	<i< br=""></i<>	<                                                                    /B                       /B             /B          /B          /B       /B																						
Price Left Delimiter		< > /B>	< > B>	< > >>TD> <td>&lt; &gt; &gt;</td> <td>&lt; &gt; </td> <td>&lt; &gt; </td> <td>&lt; &gt; /TD&gt;<td>&lt; &gt; /TD&gt;<td>&lt; &gt;  TD&gt;<td>&lt; &gt;  TD&gt;<td>&lt; &gt;  D&gt;<td>&lt; &gt;  TD&gt;&lt; &gt;  TD&gt;&lt; &gt; </td>&lt; &gt; </td>&lt; &gt; </td>&lt; &gt; </td>&lt; &gt;  &gt;&gt;TD&gt;&lt; &gt; </td>&lt; &gt;  &gt;&gt;TD&gt;&lt;   &gt;&gt;</td>	< > >	< >	< >	< > /TD> <td>&lt; &gt; /TD&gt;<td>&lt; &gt;  TD&gt;<td>&lt; &gt;  TD&gt;<td>&lt; &gt;  D&gt;<td>&lt; &gt;  TD&gt;&lt; &gt;  TD&gt;&lt; &gt; </td>&lt; &gt; </td>&lt; &gt; </td>&lt; &gt; </td>&lt; &gt;  &gt;&gt;TD&gt;&lt; &gt; </td> < >  >>TD><   >>	< > /TD> <td>&lt; &gt;  TD&gt;<td>&lt; &gt;  TD&gt;<td>&lt; &gt;  D&gt;<td>&lt; &gt;  TD&gt;&lt; &gt;  TD&gt;&lt; &gt; </td>&lt; &gt; </td>&lt; &gt; </td>&lt; &gt; </td> < >  >>TD>< >	< >  TD> <td>&lt; &gt;  TD&gt;<td>&lt; &gt;  D&gt;<td>&lt; &gt;  TD&gt;&lt; &gt;  TD&gt;&lt; &gt; </td>&lt; &gt; </td>&lt; &gt; </td> < >	< >  TD> <td>&lt; &gt;  D&gt;<td>&lt; &gt;  TD&gt;&lt; &gt;  TD&gt;&lt; &gt; </td>&lt; &gt; </td> < >	< >  D> <td>&lt; &gt;  TD&gt;&lt; &gt;  TD&gt;&lt; &gt; </td> < >	< >  TD>< >  TD>< >													

FIGURE 10A

new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No EL864389877US Attorney Docket No.: 2102680-990101 Gray Cary et al. – GTS/415-836-2500 Applicants: Victor Hsieh Sheet 8 of 36

Delimiter	Candidates
Price Right Delimiter	
	! </td
	! <
	!
	!</td
	</td
	<
Head	<
	<h< td=""></h<>
	<ht< td=""></ht<>
	<htm< td=""></htm<>
	<html< td=""></html<>
	<html></html>
Tail	>
	L>
	ML>
	TML>
	HTML>
	/HTML>

FIGURE 10B



**FIGURE 11** 

new U.S. Appln. based on Appln Mos 60/236,574 & 60/299,360 for Online Intelligent Information Express Mail Label No. EL864389877US
Attorney Docket No.: 2102680-990101
Gray Cary et al. – GTS/415-836-2500
Applicants: Victor Hsich Sheet 9 of 36

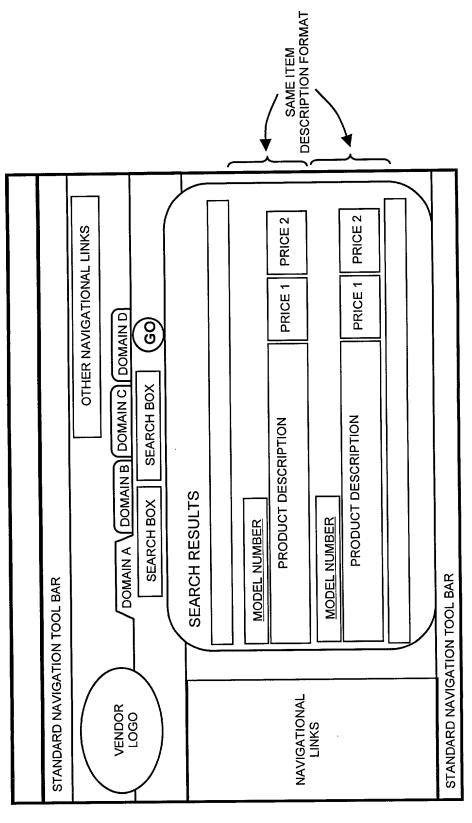


FIGURE 12

new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US
Attorney Docket No.: 2102680-990101
Gray Cary et al. – GTS/415-836-2500
Applicants: Victor Hsieh Sheet 10 of 36

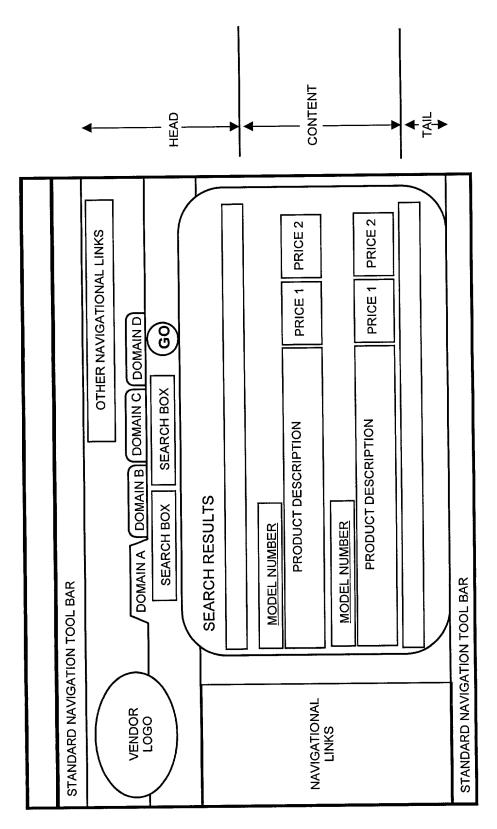
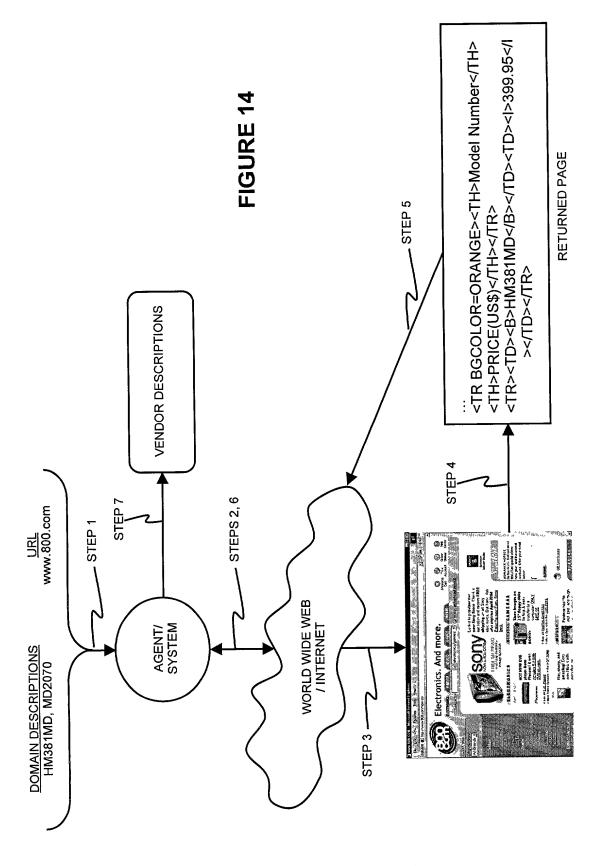


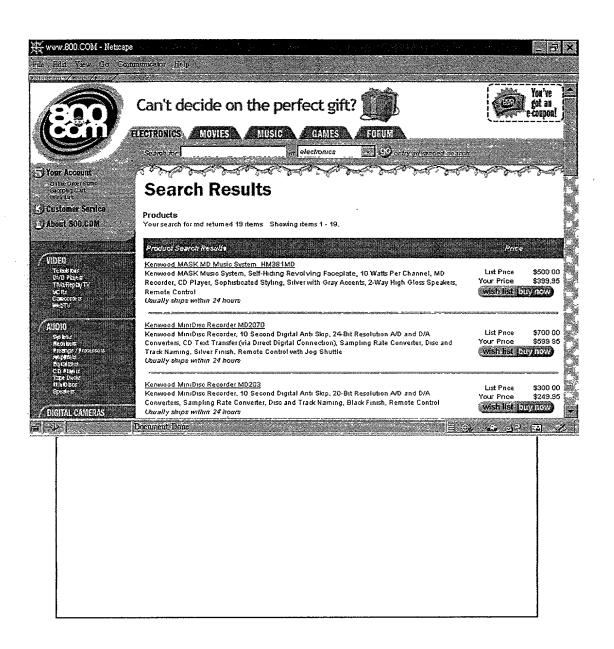
FIGURE 13

new U.S. Appln. based on Appln. Nos. 60/236,574
& 60/299,360 for Online Intelligent Information...
Express Mail Label No. EL86438987101
Attomey Docket No.: 2102680-990101
Gray Cary et al. – GTTS/415-836-2500
Applicants: Victor Hsieh Sheet 11 of 36

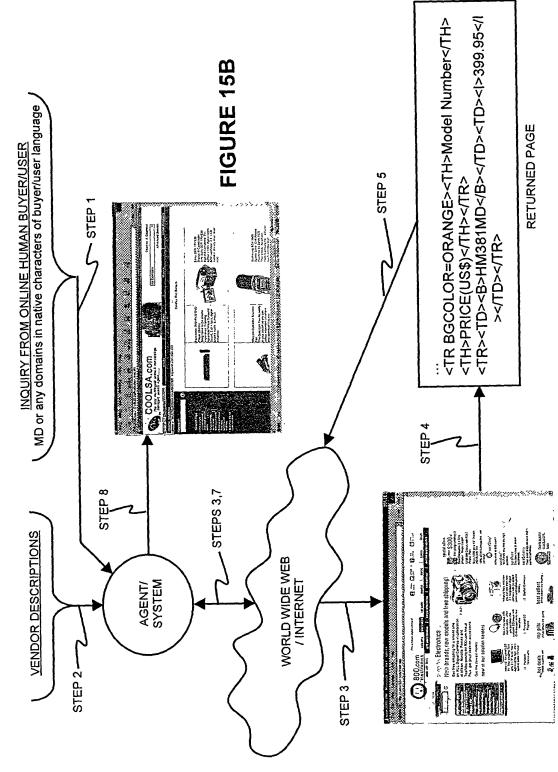


new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US
Attorney Docket No.: 2102680-990101
Gray Cary et al. – GTS/415-836-2500
Applicants: Victor Hsieh Sheet 12 of 36

new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US Attorney Docket No.: 2102680-990101 Gray Cary et al. – GTS/415-836-2500 Applicants: Victor Hsieh Sheet 13 of 36

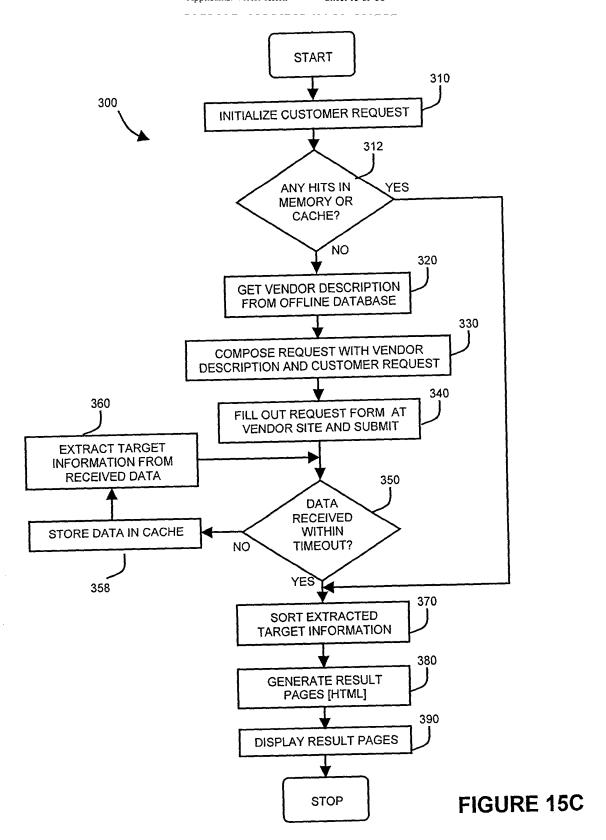


**FIGURE 15A** 



new U.S. Applin. based on Applin. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US
Attorney Docket No.: 2102680-990101
Gray Cary et al. – GTS/415-836-2500
Gray Cary et al. – GTS/415-836-2500
Applicants: Victor Hsieh Sheet 14 of 36

new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US Attorney Docket No.: 2102680-990101 Gray Cary et al. – GTS/415-836-2500 Applicants: Victor Hsieh Sheet 15 of 36



new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US Attorney Docket No.: 2102680-990101 Gray Cary et al. – GTS/415-836-2500 Applicants: Victor Hsieh Sheet 16 of 36

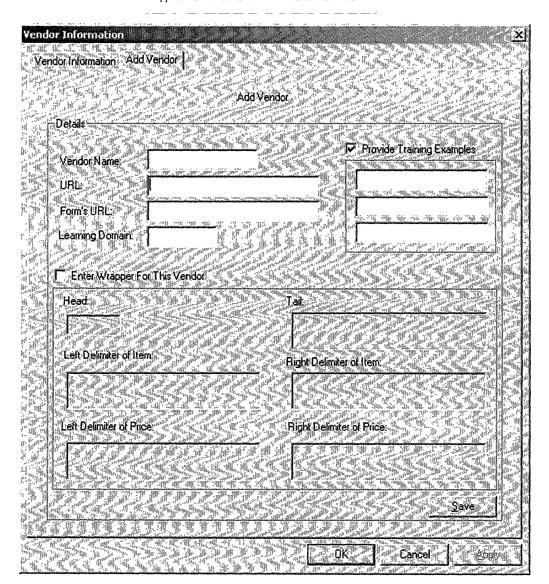


FIGURE 16

new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US Attorney Docket No.: 2102680-990101 Gray Cary et al. – GTS/415-836-2500

Applicants: Victor Hsieh Sheet 17 of 36

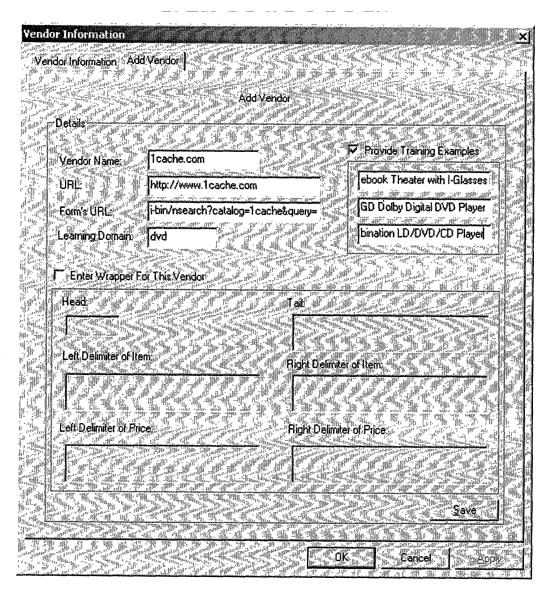


FIGURE 17

new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US Attorney Docket No.: 2102680-990101 Gray Cary et al. – GTS/415-836-2500 Applicants: Victor Hsieh Sheet 18 of 36

endor Information	Le Land Day ( Grant Strick)		
Vendor Information)   Add Vendor			
	or Information		
Vendor Name 1 1cache.com		FB	earch 1
Details			
Vendor Name Tcache com			
URL http://www.1cache.co	m	•	
	gi-bin/nsearch?catalog=1c		
Learning Domain: dvd	-groll misearch (catalog= [c		
Wapper			
Head:	Tait		
Left Delimiter of Item	I Right Delimiter of Item		
	I I I I I I I I I I I I I I I I I I I		
Left Delimiter of Price:	Right Delimiter of Price		
WARAFAR WARAN			
		<u> </u>	<u> </u>
	DK DK	Cancel	2007

FIGURE 18

new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US Attorney Docket No.: 2102680-990101 Gray Cary et al. – GTS/415-836-2500 Applicants: Victor Hsieh Sheet 19 of 36

endor Information	Aud Veridor J					**************************************
		Vendor Info	mation :			
, Vendor Name.					<u>S</u> earch	
Details						
Vendor Name:						
URL Form's URL						
Learning Domai	r.					
						, a di
-Wrapper			Taib			
Left Delimiter of	illem:		Right Delimit	er of Item		
Left Delimiter of	Price		Right Delimite	r of Price		
					<u>S</u> ave	

FIGURE 19

new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US Attorney Docket No.: 2102680-990101 Gray Cary et al. – GTS/415-836-2500

Applicants: Victor Hsieh Sheet 20 of 36

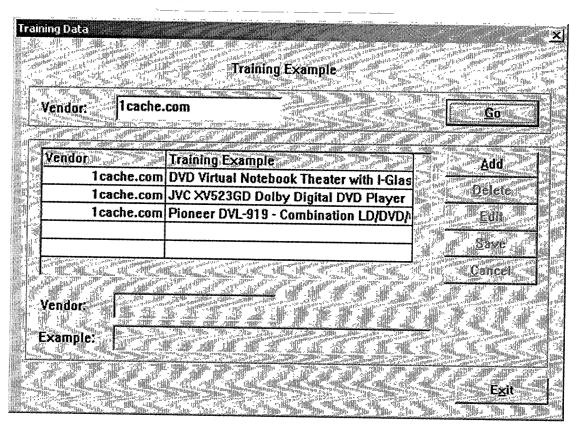


FIGURE 20

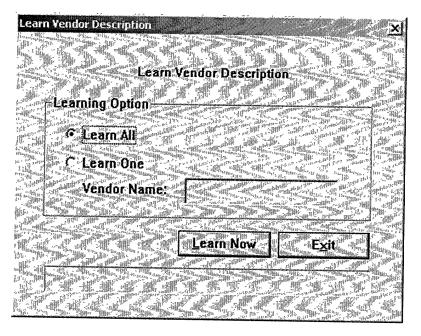


FIGURE 21

new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US Attorney Docket No.: 2102680-990101 Gray Cary et al. – GTS/415-836-2500 Applicants: Victor Hsieh Sheet 21 of 36

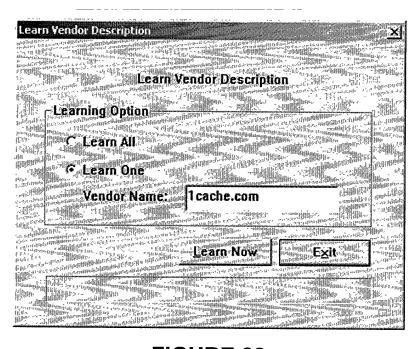


FIGURE 22

new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US Attorney Docket No.: 2102680-990101 Gray Cary et al. – GTS/415-836-2500

Applicants: Victor Hsieh Sheet 22 of 36

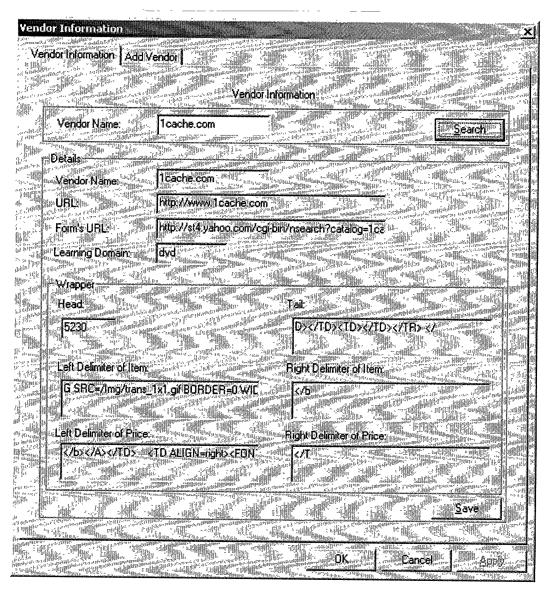


FIGURE 23

new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US Attorney Docket No.: 2102680-990101 Gray Cary et al. – GTS/415-836-2500 Applicants: Victor Hsieh Sheet 23 of 36



### FIGURE 24

```
procedure exec<sub>HLRT</sub> (wrapper (h, l<sub>i</sub>, r<sub>i</sub>, l<sub>p</sub>, r<sub>p</sub>, t), page P)

m \leftarrow 0

scan in P to line h

while the next occurrence of l_i in P occurs before the next occurrence of t

m \leftarrow m + 1

scan in P to the next occurrence of l_i in P; save position as b_{m,i}

scan in P the next occurrence of r_i; save position as e_{m,i}

scan in P to the next occurrence of l_p in P; save position as b_{m,p}

scan in P to next occurrence of r_p; save position as e_{m,p}

return label \{\ldots, << b_{mi}, e_{mi}>, < b_{mp}, e_{mp}>> \}
```

### FIGURE 25

### Procedure learnHLRT (examples ε)

- 1. Generate the candidate sets  $Cands_r(I, \varepsilon)$ ,  $Cands_l(I, \varepsilon)$ ,  $Cands_l(p, \varepsilon)$ ,  $Cands_l(p, \varepsilon)$ .
- 2. Enumerate the cross product of these candidate sets; each element W = (h,  $l_i$ ,  $r_i$ ,  $l_p$ ,  $r_p$ , t), of this cross product is a wrapper. Halt if W is satisfactory i.e., execLR(W,  $P_n$ ) = Ln, for every ( $P_n$ , Ln)  $\in \epsilon$ .

## FIGURE 26

new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US Attorney Docket No.: 2102680-990101 Gray Cary et al. – GTS/415-836-2500 Applicants: Victor Hsieh Sheet 24 of 36

```
* The main problem
 */
 procedure learnHLRT (examples ε) for each 1 <k <K
 for each u \in cands_i(p, \epsilon): if valid_i(u, p, \epsilon) then I_p \leftarrow u and terminate this loop
 for each u \in cands_r(i, \epsilon): if valid_r(u, i, \epsilon) then r_i \leftarrow and terminate this loop
 for each u \in cands_r(p, \varepsilon): if valid_r(u, p, \varepsilon) then r_p \leftarrow u and terminate this loop
 for each u_{ii} \in cands_i(i, \varepsilon)
 for each u_h \in cands_h(\varepsilon)
 for each u_t \in cands_t(\epsilon)
 if valid<sub>ih.t</sub>(u<sub>li</sub>, u<sub>h</sub>, u<sub>t</sub>, ε) then
 I<sub>i</sub>←u<sub>ii</sub>, h←u<sub>h</sub>, t←u<sub>t</sub>, and terminate these three loops
 return HLRT wrapper (h, l<sub>i</sub>, r<sub>i</sub>, l<sub>p</sub>, r<sub>p</sub>, t)
 *Generate a set of candidates for left delimiter of price attribute
procedure cands<sub>i</sub>(attribute price, examples ε)
 return the set of all suffixes of the shortest string in neighbors<sub>i</sub>(price, ε)
 *Generate a set of candidates for right delimiter of attribute a. Here a could be item or price
procedure cands<sub>r</sub>(attribute a, examples ε)
 return the set of all prefixes of the shortest string in neighbors, (a, \varepsilon)
 * Generate a set of candidates for a page's head
procedure cands<sub>h</sub> (examples ε)
 return the set of all substrings of the shortest string in heads(E)
 *Generate a set of candidates for a page's tail
procedure cands, (example \varepsilon)
 return the set of all substrings of the shortest string in tails(ε)
 *Determine whether a particular candidate for the left delimiter of price is valid
 *This procedure applies constraints C3
 procedure valid(candidate u, attribute price, examples E)
 for each s∈ neighbors₁(price, ε): if u is not a proper suffix of s then return FALSE
```

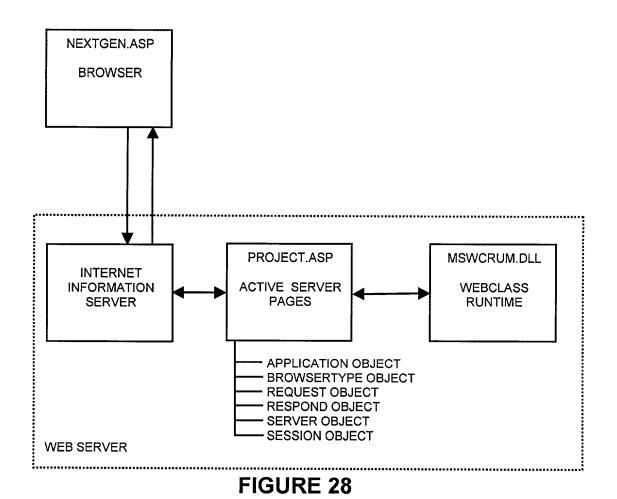
Figure 27A

new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US Attorney Docket No.: 2102680-990101 Gray Cary et al. – GTS/415-836-2500 Applicants: Victor Hsieh Sheet 25 of 36

```
for each s \in tails(\epsilon): If u is a sub string of s then return FALSE
 return TRUE
 *Determine whether a particular candidate for the right delimiter of item or price is
 *This procedure applies constraints C1 and C2
 procedure valid<sub>r</sub>(candidate u, attribute a, examples ε)
for each s ∈ attribs(a, ε): if u is a sub string of s then return FALSE
for each s \in neighbors(a, \epsilon): if u is not a prefix of s then return FALSE
 return TRUE
 *Determine whether a particular combination of candidate uh, ut, and uli for h, t, and
 *li respectively are satisfactory.
procedure valid<sub>ii,h,t</sub>(candidates u_{ii}, u_h, u_t, example \epsilon)
 for each s \in heads(\varepsilon)
 if is not a sub string of s then return FALSE
 if Un is not a proper suffix of scan (s, Uh) then return FALSE
 if Ut occurs before U11 in scan (s, Uh), then return FALSE
 for each s \in tails (\epsilon)
 if Ut 1S not a sub string of s then return FALSE
 if Uli occurs before Ut in s then return FALSE
 for each s \in seps(\epsilon)
 if Uli is not a proper suffix of s then return FALSE
 if Ut occurs before Uli in s then return FALSE
 return TRUE
*Return a set of containing all values of a attribute, either item(i) or price(p) in each
*example
 procedure attribs(attribute a, examples ε)
 return U_{(Pn,Ln)} \in \varepsilon(Pn[b_{m,a}, e_{m,a}] \mid <..., (b_{m,k}, e_{m,k}), ... > \in Ln
*Return all strings to the left of an attribute, whether these strings are in the heads
* or the bodies of the pages
 procedure neighbors<sub>i</sub>(attribute a, examples ε)
 if a=i then return seps (i, \varepsilon) U heads (\varepsilon) else return seps (a, \varepsilon)
```

FIGURE 27B

new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US Attorney Docket No.: 2102680-990101 Gray Cary et al. – GTS/415-836-2500 Applicants: Victor Hsieh Sheet 26 of 36



CLIENT

ICONNECTION POINT

LINK

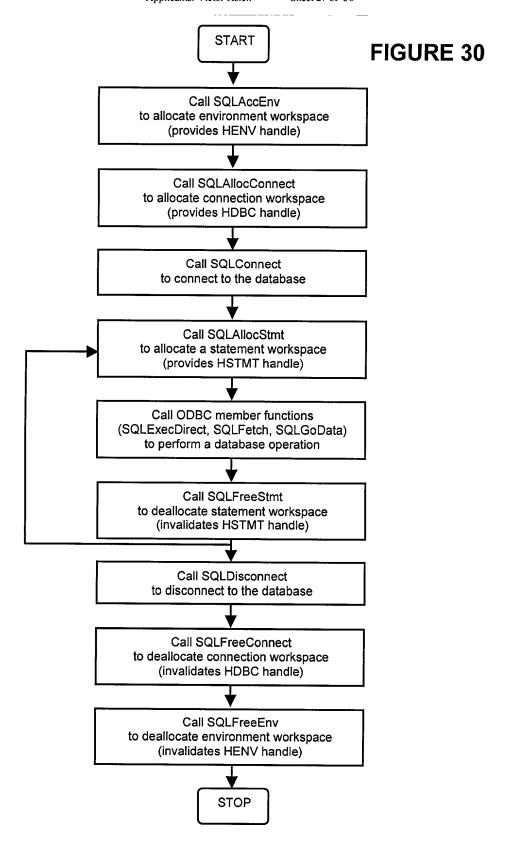
O CONNECTION POINT

CONNECTION POINT

OUTGOING INTERFACE

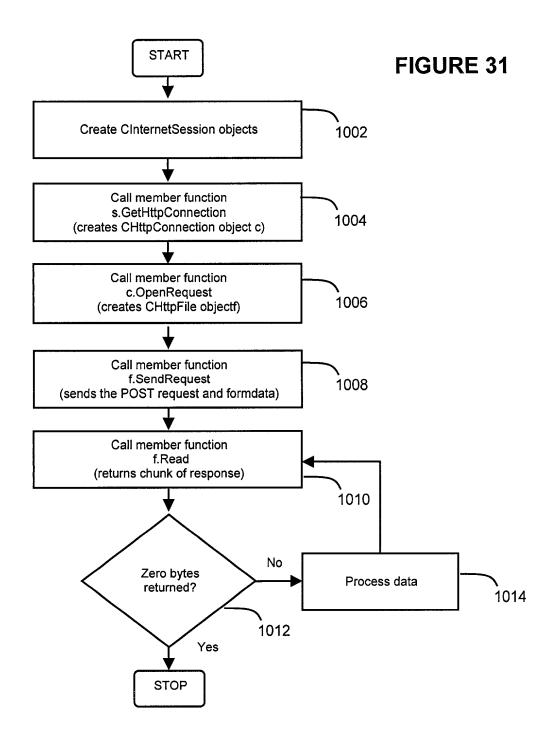
FIGURE 29

new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US Attorney Docket No.: 2102680-990101 Gray Cary et al. – GTS/415-836-2500 Applicants: Victor Hsieh Sheet 27 of 36



new U.S. Appin. based on Appin. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US Attorney Docket No.: 2102680-990101 Gray Cary et al. – GTS/415-836-2500

Applicants: Victor Hsieh Sheet 28 of 36



	NAVIGATION BUTTONS TO OTHER PAGES IN WEBSITE	
WEBSITE LOGO	C All Ontological Domains	Choose a Channel
QUICK SEARCH		Advanced Search
Enter a "MULTILIGUAL CHARACTER" keyword and click GO!	CTER" keyword and click GO!	
Keyword: Goi		
* MEMBER SIGN IN	WEBSITE INTRODUCTORY TEXT AND IMAGES	Reverse Advertising Sites
Login Reset		
ed Vendor Location 23ezid	<b>(</b> )Feedback	
pkkw.com er5lcin.com	Submit	

new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US
Attorney Docket No.: 2102680-990101
Gray Cary et al. – GTS/415-836-2500
Applicants: Victor Hsieh Sheet 29 of 36

FIGURE 32

	NAVIGATION BI	NAVIGATION BUTTONS TO OTHER PAGES IN WEBSITE	GES IN WEBSITE	
WEBSITE	0907		Choose a Channe	Choose a Channel MENT DOMAINS
				Advanced Search
DOMAIN A	DOMAIN B	DOMAIN C	DOMAIN D	OTHER DOMAINS
QUICK SEARCH				
Enter a "MULTILIGU,	Enter a "MULTILIGUAL CHARACTER" keyword and click GO!	and click GO!		
Keyword:	Company A Infomation about G2B Company A	ompany A		
SEARCH TIPS *	Company G Information about G2B Company G	Company G	<u></u>	Reverse Advertising Sites
Registered Vendor Links Abcd.com cdef.com 1234z.com 23ezidcom	Links Preedback	<b>Y</b>		
. 1	Submit			

# FIGURE 33

new U.S. Applin. based on Applin. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US
Attorney Docket No.: 2102680-990101
Gray Cary et al. — GTS/415-836-2500
Applicants: Victor Hsieh Sheet 30 of 36

	NAVIGATIO	N BUTTONS TO	NAVIGATION BUTTONS TO OTHER PAGES IN WEBSITE	S IN WEBSITE	
WEBSITE	0907		IR	Choose a Cha	Choose a Channel rement Domains
			_	Adv	Advanced Search
DOMAIN A	DOMAIN B	loa	DOMAIN C	DOMAIN D	OTHER DOMAINS
ADVANCED SEARCH	CH AGENTS ARE ON!				
Enter a "MULTILIGL	Enter a "MULTILIGUAL CHARACTER" keyword and click GO!	word and click G	iOi		
	Select All	Multilin	Multilingual Government Networks	Networks	
	Public Body	Speed	Air/Shipping/G	Air/Shipping/Ground/Free Overnight Del.	Ease of Navigation
00 F 00 4 0	☐G2B Company G	MMM		î	EEEE
* * * * * * * * * * * * * * * * * * *					
•					
	Time Out				
	Reserved Price Range	e From \$			
	Manufacturer or OEM			ı	
	Keyword:			(ÎO)	

new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US
Attorney Docket No.: 2102680-990101
Gray Cary et al. – GTS/415-836-2500
Applicants: Victor Hsieh Sheet 31 of 36

FIGURE 34

	NAVIGATION BUTTONS TO OTHER PAGES IN WEBSITE	THER PAGES IN WEBSITE	
WEBSITEL	090	C All Ontological Domains	Choose a Channel mains
	)		Advanced Search
DOMAIN A	DOMAIN B DOMAIN C	IN C DOMAIN D	D OTHER DOMAINS
QUICK SEARCH			
Enter a "MULTILIGUAL	AL CHARACTER" keyword and click GO!		
Keyword:  SEARCH TIPS  *	For B2B Agricultural / Farming Trading Exchanges, Apparel Auctions, Vertical Automotive/ Transportation Portals, Production Business Machines and Non-Production e- Computers & Electronics Procurement Supplies Industrial Equipment Medical & Scientivic Equipment	<u> </u>	Reverse Advertising Sites
		ı ravei & Enterainment	
Registered Vendor Links Abcd.com cdef.com 1234z.com 23ezidcom	Links Ding Feedback Ding Company Compa		
DKKW.COIII EIDICIII.	Submit		

FIGURE 35

new U.S. Appln. based on Appln. Mos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US
Attorney Docket No.: 2102680-990101
Gray Cary et al. – GTS/415-836-2500
Applicants: Victor Hsieh Sheet 32 of 36

	NAVIGATIC	NAVIGATION BUTTONS TO OTHER PAGES IN WEBSITE	O OTHER PAGE	ES IN WEBSITE	
				Choo	Choose a Channel
WEBSITE	LOGO			All Ontological Domains	D
				Adve	Advanced Search
DOMAIN A	DOMAIN B	ioa	DOMAIN C	DOMAIN D	OTHER DOMAINS
ADVANCED SEARC	ADVANCED SEARCH AGENTS ARE ON!				
Enter a "MULTILIGL	Enter a "MULTILIGUAL CHARACTER" keyword and click GO!	word and click G	io		
	Select All	Multiling	Multilingual B2B Suppliers	iers	
	Online Stores	Speed	Air/Shipping/(	Air/Shipping/Ground/Free Overnight Del.	Ease of Navigation
	☐B2B Company 1	MMM		Ϋ́	EEEE
SEARCH TIPS  *	☐ B2B Company 2	fifffff		ſſ	EEEEE
* +	☐ B2B Company 3	MMM		Ϋ́	EEEE
	☐ B2B Company 4	MMM		Λ	EEEE
	Time Out				
	Price Range	From \$	To \$	φ.	
	Manufacturer			(	
	Keyword:			(coi)	

**FIGURE 36** 

new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information .. Express Mail Label No. EL864389877US
Attorney Docket No.: 2102680-990101
Gray Cary et al. – GTS/415-836-2500
Applicants: Victor Hsieh Sheet 33 of 36

	NAVIGAT	NAVIGATION BUTTONS TO OTHER PAGES IN WEBSITE	ER PAGES IN W	EBSITE
				Choose a Channel
WEBSITELO	090			All Domains
	) )			Advanced Search
DOMAIN A	DOMAIN B	DOMAINC		DOMAIN D OTHER DOMAINS
QUICK SEARCH				
Enter a "MULTILIGUAL (	HARACTER" ke	L CHARACTER" keyword and click GO!		
Keyword:		Hot Sellers Domain A	omain A	
SEARCH TIPS	Image Item A	Description Item A	Image Item B	Description Item B
* * *	Image Item	Description Item 1	Image Item G	Description Item G
	Image Item A	Description Item A	Image Item K	Description Item K
Registered Vendor Links  Abcd.com cdef.com		<b>▶</b> Feedback		. [
1234z.com 23ezidcom pkkw.com er5lcin.com	n Submit			

Attorney Docket No.: 2102680-990101 Gray Cary et al. – GTS/415-836-2500 Applicants: Victor Hsieh Sheet

new U.S. Applin. based on Applin. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL864389877US

Sheet 34 of 36

FIGURE 37

	NAVIGATIO	T SŅOTTUB NC	O OTHER PAG	NAVIGATION BUTTONS TO OTHER PAGES IN WEBSITE	
				Choo	Choose a Channel
WEBSITE	LOGO			All Domains	D
	:		]	Adva	Advanced Search
DOMAIN A	DOMAIN B	00	DOMAIN C	DOMAIN D	OTHER DOMAINS
ADVANCED SEAR	ADVANCED SEARCH AGENTS ARE ON!				
Enter a "MULTILIGE	Enter a "MULTILIGUAL CHARACTER" keyword and click GO!	word and click G	iō		
	Select All		Engli	English Vendors	
	Online Stores	Speed	Air/Shipping/C	Air/Shipping/Ground/Free Overnight Del.	Ease of Navigation
1	☐ Store 1AAC	IIII		1	EEEE
SEARCH TIPS	☐ Store 1443A	MMMM		ſΥ	EE
* +	☐ Store 1443A	MMM		ſſ	EE
	Select All		GE!	Chinese Vendors	
	Store C5A3	MMMM		Î	EEEE
	☐ Store DRGWS4	በበበበበ		Î	EEEE
	Time Out				
	Price Range	From \$		69	
	Manufacturer				
	Keyword:			(60)	

new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. E102680-990101 Attomey Docket No.: 2102680-990101 Gray Cary et al. – GTS/415-836-2500 Applicants: Victor Hsieh Sheet 35 of 36

**-IGURE 38** 

	NAVIGATION B	NAVIGATION BUTTONS TO OTHER PAGES IN WEBSITE	GES IN WEBSITE	
			Ö	Choose a Channel
WEBSITE I	-060		ALL DOMAINS	D
			A	Advanced Search
DOMAIN A	DOMAIN B	DOMAIN C	DOMAIN D	OTHER DOMAINS
QUICK SEARCH				
Enter a "MULTILIGUAI	AL CHARACTER" keyword and click GO!	and click GO!		
Keyword:	Search Results			
	Products Your search for XX returned 5 Domain C items.	rned 5 Domain C items.		
	Model / Features	es	Store	Price
* •	AMS7 XX Widget [Features]		Store 1443A	\$399.95 Buy now
	AMS2 XX Widget [Features]		Store 1443A	\$329.95 Buy now
	AMS7 XX Widget [Features]		Store DRGWS4	\$459.95 Buy now
	AMS7 XX Widget [Features]		Store 4U\$U\$	\$499.95 Buy now
	ZRTQ XX Widget [Features]		Store OOHE	\$529.95 Buy now

**FIGURE 39** 

new U.S. Appln. based on Appln. Nos. 60/236,574 & 60/299,360 for Online Intelligent Information... Express Mail Label No. EL86498987101 Attorney Docket No.: 2102680-990101 Gray Cary et al. – GTS/415-836-2500 Gray Cary et al. – GTS/415-836-2500